

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 1 132 896 A1

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:
12.09.2001 Bulletin 2001/37

(51) Int Cl.7: **G10L 21/02**(21) Application number: **00400623.5**(22) Date of filing: **08.03.2000**

(84) Designated Contracting States:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE**
Designated Extension States:
AL LT LV MK RO SI

(72) Inventor: **Bourzeix, François**
31000 Toulouse (FR)

(74) Representative: **David, Alain et al**
Cabinet Beau de Loménie
158, rue de l'Université
75340 Paris Cedex 07 (FR)

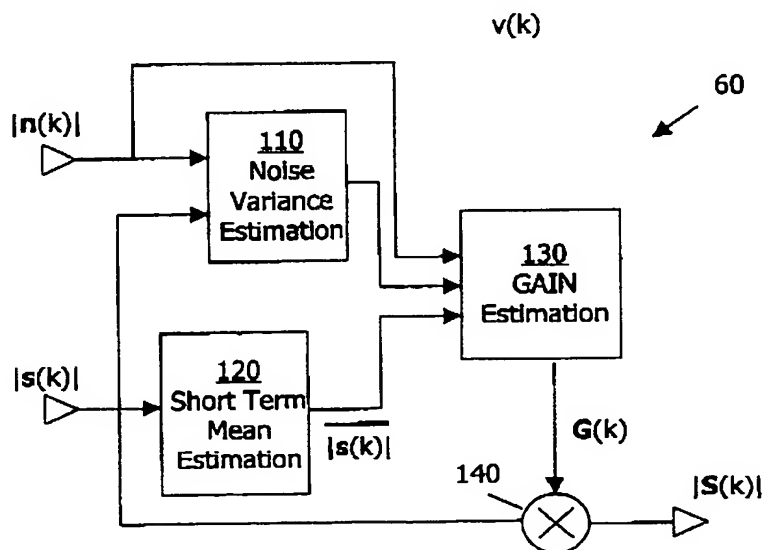
(71) Applicant: **MOTOROLA, INC.**
Schaumburg, IL 60196 (US)

(54) **Frequency filtering method using a Wiener filter applied to noise reduction of acoustic signals**

(57) A method and apparatus for suppressing acoustic noise in an acoustic signal ($s(n)$) represented by the frequency components of a plurality of frames each representing a small portion of the acoustic signal, comprising estimating the average magnitude of noise in each frequency component over a plurality of frames; estimating the variability of the magnitude of noise (110) in each frequency component; and generating denoising filter components in dependence on the estimated noise magnitudes, the estimated variability of the noise

magnitude in each frequency component and the magnitude of each frequency component, and varying the magnitude of each frequency component (140) in dependence on the corresponding denoising filter component.

This has the significant advantage of taking account of the variability of the magnitude of noise within each frequency component over time, making possible determination of an approximate probability of any one frequency component being largely comprised of noise or alternatively of wanted speech signal.

FIG. 2

Description

Field of the Invention

[0001] The present invention relates to a method and apparatus for processing an acoustic signal and in particular for suppressing background acoustic noise from the acoustic signal.

Background of the Invention

[0002] Portable communication devices such as cellular telephones often need to detect and transmit a speech or similar signal in noisy environments such as a fast moving vehicle. Methods of suppressing background acoustic noise have thus been developed which permit much better communication with such devices in noisy environments.

[0003] The general approach adopted by some such methods is to represent the overall acoustic signal as the frequency components of a plurality of frames, each frame basically representing a small portion (e.g., about 10 ms) of the acoustic signal, and then to attempt to detect and remove or suppress any noise components occurring within each frequency component of each frame.

[0004] One simple and crude method estimates the average magnitude of noise $|n(k)|$ in each frequency component $s(k)$ over a large number of frames (e.g., of the order of about 100 frames) and then simply subtracts this from the magnitude of the frequency component $|s(k)|$ to generate modified or noise suppressed frequency component magnitudes $|S(k)|$. This method can be enhanced by providing that the magnitudes $|S(k)|$ of the modified frequency components are never allowed to fall below a minimum comfort noise floor level η .

[0005] A more sophisticated method, known as wiener filtering, multiplies the magnitudes $|s(k)|$ of the frequency components in each frame by a denoising filter having components $G(k)$ such that $|S(k)| = G(k) \cdot |s(k)|$, where the $G(k)$ are generated in respect of each frame according to the formula: -

$$G(k) = \frac{E\{|X(k)|^2\}}{E\{|X(k)|^2\} + |n(k)|^2}$$

where

$$E\{|X(k)|^2\}$$

is the expected value of the square of the magnitude of the clean or denoised speech. Clearly, in a real system this must be estimated (e.g., by assuming that $X(k) \approx S(k)$).

[0006] In a third method, known as Minimum Mean Square Estimation (MMSE), the magnitudes $|s(k)|$ of the frequency components $s(k)$ are again multiplied by denoising filter components $G(k)$ such that $|S(k)| = G(k) \cdot |s(k)|$, but in this case, the $G(k)$ are estimated using modified Bessel functions (which must be sampled). This method is intensive in the amount of processing power which it requires (in terms of Millions of Instructions Per Second (MIPS)) which makes it unsuitable for portable communication devices where processing power is at a premium.

[0007] Furthermore, none of the above methods takes any account of the variability of the noise components over time.

Summary of the Invention

[0008] According to a first aspect of the present invention, there is provided a method of suppressing acoustic noise in an acoustic signal represented by the frequency components of a plurality of frames, each frame representing a small portion of the acoustic signal, comprising the steps of estimating the average magnitude of noise in each frequency component over a plurality of frames, estimating the variability of the magnitude of noise in each frequency component; and generating denoising filter components in dependence on the estimated magnitude of noise in each frequency component, the estimated variability of the magnitude of noise in each frequency component and the magnitude of each frequency component, and varying the magnitude of each frequency component in dependence on the corresponding denoising filter component.

[0009] This method has the significant advantage of taking account of the variability of the magnitude of noise within each frequency component over time. In this way, it is possible to determine an approximate probability of any one frequency component being largely comprised of noise or alternatively of being largely comprised of wanted speech signal.

[0010] Preferably, the method further comprises setting the filter components in dependence on the ratio of the magnitude of each frequency component to an estimated likely maximum magnitude of noise for that frequency component, whereby, if the ratio exceeds a predetermined amount for a given frequency component, the filter component corresponding to such a frequency component may be set to a maximum value which is preferably substantially equal to one, whereas, if the ratio is less than a second predetermined amount, the corresponding filter component may be set to a minimum value, which is preferably substantially equal to 0.15. In one preferred embodiment, the filter components are varied in a linear dependence on the ratio of the magnitude of each frequency component to the estimated likely maximum magnitude of noise for that frequency component between a minimum value of the filter components at or below the second predetermined amount

and a maximum value at or above the first predetermined value.

[0011] By having a maximum value of each filtering component of about 1, it provides that for signals which are much larger than the maximum likely noise content, there is no signal attenuation. This is actually very beneficial for frequency components which are much larger than the likely maximum noise component, because since the phase of the noise component is not necessarily aligned to the phase of the speech (or similar) signal, the noise component is almost as likely to destructively interfere with the speech signal (thus reducing the overall magnitude of the frequency component relative to the clean speech equivalent) as to constructively interfere with it (thus increasing the magnitude of the frequency component relative to the clean speech equivalent).

[0012] According to a further preferred embodiment of the present invention, prior to calculating the ratio of the magnitude of each frequency component to the estimated likely maximum magnitude of noise for that frequency component, the magnitudes of the frequency components are filtered to remove high frequency fluctuations thereof. This filtering of the magnitudes of the frequency components is preferably achieved by generating a short term mean estimation of the mean magnitudes of the frequency components, preferably over approximately three frames.

[0013] According to a second aspect of the present invention, there is provided an apparatus for suppressing acoustic noise in an acoustic signal represented by the frequency components of a plurality of frames, each frame representing a small portion of the acoustic signal, comprising: means for estimating the average magnitude of noise in each frequency component over a plurality of frames; means for estimating the variability of the magnitude of noise in each frequency component; means for generating denoising filter components in dependence on the estimated magnitude of noise in each frequency component, the estimated variability of the magnitude of noise in each frequency component and the magnitude of each frequency component, and means for varying the magnitude of each frequency component in dependence on the corresponding denoising filter component.

Brief Description of the Drawings

[0014] In order that the present invention may be better understood, embodiments thereof will now be described, by way of example only, with reference to the accompanying drawings in which:-

FIG. 1 is a block diagram of apparatus or method steps suitable for carrying out the present invention; and

FIG. 2 is a more detailed block diagram of noise

suppression apparatus or method steps in the apparatus or method of FIG. 1.

Detailed Description of the Invention

[0015] Referring firstly to FIG. 1, there is shown a series of steps or apparatus blocks showing the overall approach of noise suppression according to the present invention. Considering FIG. 1 initially as representing a series of method steps, these are now described in detail below.

[0016] The first step 10 is to take the acoustic signal $s(n)$ (in the form of digital audio signal amplitude samples) and to perform high pass filtering to remove low frequency components (which do not carry much speech signal information although they may contain a large amount of unwanted background acoustic noise).

[0017] The second step 20 windows and overlaps (for example, by 50%) the high pass filtered acoustic signal. This step involves separating the signal into a series of overlapping segments and windowing them to form frames so that at the edge of each frame the amplitude of the signal is zero.

[0018] The third step 30 performs the Fast Fourier Transform on each windowed vector. Given a 256 input signal vector $s(n)$, we obtain a 256 vector $s(k)$ where n and k stand respectively for some time, and frequency indices. In what follows we shall indicate spectral data with bold characters: n, s, \dots

[0019] The fourth step 40 performs a transformation of the FFT outputs, from Cartesian to polar co-ordinates.

[0020] The fifth step 50 uses the magnitude of the Fourier Transform, to evaluate the mean magnitude of spectral background noise $\text{mag}(n(k))$.

[0021] The sixth step 60 performs the estimation of de-noised speech spectral magnitude $\text{mag}(s(k))$ using the noise evaluation from block 50, and the noisy speech spectral magnitude.

[0022] The seventh, eighth and ninth steps (70,80,90) perform the symmetrical operations to those performed by respectively 30,20 and 10: conversion from polar to Cartesian, inverse Fourier transforms and overlap add. It is to be noted that the signal phases is not modified by the algorithm since the noisy speech phases is used to reconstruct the clean speech signal in step 70.

The main structure of this algorithm is very classical. The innovative feature of the algorithm is in the way noise is removed from speech in step 60. This step is now described in detail.

[0023] Referring now to FIG. 2, the step 60 can be subdivided into 3 sub-steps.

[0024] The first sub-step 110 is dedicated to evaluating the noise variance. Step 50 output is the mean magnitude of background noise. Thus on speechless frames 55 input data can be used to evaluate the noise variance

$$\sigma(k) = \text{mean}(\text{mag}(s(k)-n(k)))/\text{mag}(n(k)).$$

In fact the variance is obtained by low filtering the selected speech-free $s(k)$:

$$\sigma^p(k) = \delta \cdot \sigma^{p-1}(k) + (1-\delta) \text{mag}(s(k)-n(k))/\text{mag}(n(k)) \quad 5$$

where superscript p indicates the number of the speechless frame. For a given frequency channel, a sample is considered to be only noise if $\text{mag}(s(k)) < 4 \cdot \text{mag}(n(k))$.

[0025] The second sub-step 120 is dedicated to evaluating the input signal short term mean $M(k)$ (smoothed version of $s(k)$). It is obtained thanks to a one-tap IR filter:

$$M^q(k) = \gamma \cdot M^{q-1}(k) + (1-\gamma) \text{mag}(s(k)) \quad 10$$

where superscript q indicates the frame number.

[0026] The third sub-step 130 is dedicated to calculating the denoising filter gain for each frequency channel. It is done as follows:

- If $M(k)/n(k) < 1$ then it is considered that there is only noise and the minimum gain factor K is applied, thus $G(k)=K$.
- If $M(k)/n(k) > 1 + \beta\sigma(k)$ it is considered that noise is negligible and $G(k)$ is set to 1, which is the maximum gain factor (β is typically equal to 20).
- In between we use a linear interpolation to calculate the gain factor: $G(k) = K + (1-K) (M(k)/n(k) - 1) / (\beta\sigma(k))$.

[0027] The last operation of the algorithm consists in applying, at mixer 140, the gain to the noisy speech spectral magnitude to obtain the estimation of the clean speech: $\text{mag}(S(k)) = G(k) \cdot \text{mag}(s(k))$.

Claims

1. A method of suppressing acoustic noise in an acoustic signal ($s(k)$) represented by the frequency components of a plurality of frames, each frame representing a small portion of the acoustic signal, comprising the steps of estimating the average magnitude of noise (50) in each frequency component over a plurality of frames, estimating the variability of the magnitude of noise (110) in each frequency component; and generating de-noising filter components (60) in dependence on the estimated magnitude of noise in each frequency component, the estimated variability of the magnitude of noise in each frequency component and the magnitude of each frequency component, and varying the magnitude of each frequency component (140) in dependence on the corresponding de-noising filter component.
2. The method according to claim 1 further comprising

calculating the ratio of the magnitude of each frequency component to an estimated likely maximum magnitude of noise for that frequency component and setting the filter components in dependence on the calculated ratio for that frequency component, whereby, if the ratio exceeds a predetermined amount, which depends on the noise estimated variability of the magnitude, for a given frequency component, the filter component corresponding to such a frequency component may be set to a maximum value, whereas, if the ratio is less than a second predetermined amount, the corresponding filter component may be set to a minimum value.

3. The method according to claim 2 wherein the filter components are varied in a linear dependence on the ratio of the magnitude of each frequency component to the estimated likely maximum magnitude of noise for that frequency component between a minimum value of the filter components at or below the second predetermined amount and a maximum value at or above the first predetermined value.
4. The method according to claim 2 or 3 wherein the first predetermined value is in a linear dependence to the estimated variability of the noise variability.
5. The method according to claim 2 or 3 wherein the minimum value is substantially equal to 0.15.
6. The method according to claim 4 wherein the variability of the noise magnitude for a given frequency component is estimated by filtering on speechless frames, with a one tap IR filter, the distance between the noise estimated magnitude and the noisy speech magnitude, divided by the noise estimated magnitude.
7. The method according to claim 2 wherein prior to calculating the ratio of the magnitude of each frequency component to the estimated likely maximum magnitude of noise for that frequency component, the magnitudes of the frequency components are filtered to remove high frequency fluctuations thereof.
8. The method according to claim 6 wherein the filtering of the magnitudes of the frequency components is achieved by generating a short term mean estimation of the mean magnitudes of the frequency components.
9. An apparatus for suppressing acoustic noise in an acoustic signal ($s(k)$) represented by the frequency components of a plurality of frames, each frame representing a small portion of the acoustic signal, comprising:

means for estimating the average magnitude of noise in each frequency component over a plurality of frames (50);

means for estimating the variability of the magnitude of noise in each frequency component (110);

means for generating de-noising filter components (60) in dependence on the estimated magnitude of noise in each frequency component, the estimated variability of the magnitude of noise in each frequency component and the magnitude of each frequency component, and

means for varying the magnitude of each frequency component (140) in dependence on the corresponding de-noising filter component.

20

25

30

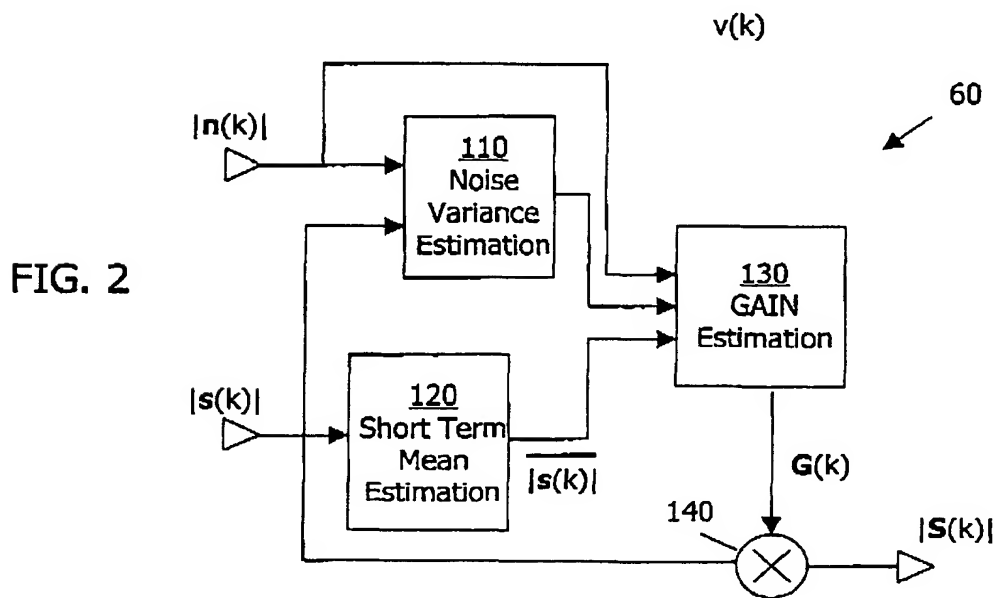
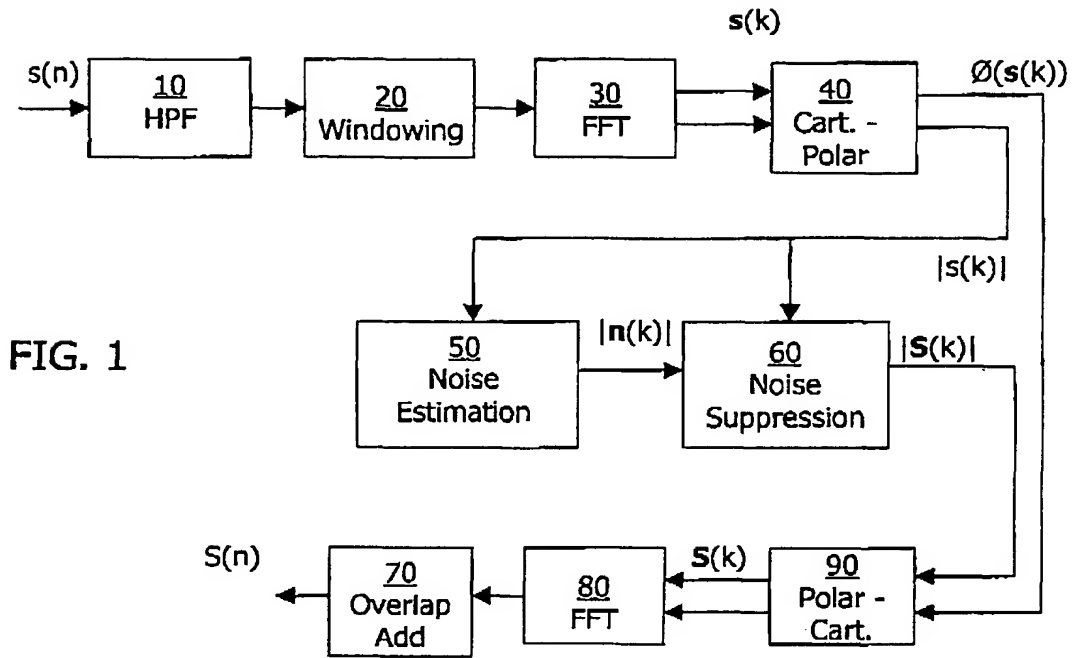
35

40

45

50

55





European Patent
Office

EUROPEAN SEARCH REPORT

Application Number
EP 00 40 0623

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
X	EP 0 918 317 A (SEXTANT AVIONIQUE) 26 May 1999 (1999-05-26) * abstract * * Equations 8, 9 *	1-3, 7-9	G10L21/02
X	FEI X ET AL: "Speech enhancement by spectral magnitude estimation - A unifying approach" SPEECH COMMUNICATION, NL, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, vol. 19, no. 2, 1 August 1996 (1996-08-01), pages 89-104, XP004013500 ISSN: 0167-6393 * figure 5 *	1, 9	
A	EP 0 913 810 A (SONY CORP) 6 May 1999 (1999-05-06) * page 5, line 40 - page 7, line 37 *	1, 9	
			TECHNICAL FIELDS SEARCHED (Int.Cl.7)
			G10L
The present search report has been drawn up for all claims			
Place of search THE HAGUE		Date of completion of the search 31 July 2000	Examiner Krembel, L
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document			

EPO FORM 1503 (3.92 (Rev.01))

**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**

EP 00 40 0623

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

31-07-2000

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0918317 A	26-05-1999	FR 2771542 A	28-05-1999
		JP 11265198 A	28-09-1999
EP 0913810 A	06-05-1999	JP 11133992 A	21-05-1999
		AU 8937398 A	20-05-1999
		BR 9804324 A	21-12-1999
		CN 1216380 A	12-05-1999

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82